DOCUMENT RESUME

ED 302 564 TM 012 564

AUTHOR Stanley, Sarah J.
TITLE A Convergence or Divergence? The Evolving Evaluation Advice of Scriven, Stake, and Patton.
PUB DATE 87
NOTE 29p.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Attitudes; *Educational Assessment; Educational Philosophy; Elementary Secondary Education; Evaluation; *Evaluation Utilization; Evaluators; *Theories
IDENTIFIERS *Theoretical Orientation

ABSTRACT

This paper considers the manners in which three evaluation theorists may have altered their views regarding two major evaluation-related issues under consideration during the past 20 years. The theorists are Michael Scriven, Robert Stake, and Michael Patton, who are characterized as the philosopher-evaluator, responsive-evaluator, and utilization-focused evaluator, respectively. The issues under consideration concern: (1) the extent to which results of educational evaluations should be directed to single, as opposed to multiple, audiences; and (2) the degree to which educational evaluations should be attentive to the likelihood that the results will be used by decision-makers. Although at the outset of their influence on the evaluation community their positions on the two questions could be easily distinguished, the theorists' views have since converged toward the same goals. Mutual influences are apparent in the development of the theorists' views on both issues. (TJH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED302564

A CONVERGENCE OR DIVERGENCE?

The Evolving Evaluation Advice of Scriven, Stake, and Patton

by

Sarah J. Stanley

of the

University of California, Los Angeles

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

SARAH J. STANLEY

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 012 564

BEST COPY AVAILABLE

Most of us profit from experience. Although there are surely a few exceptions, educational evaluators have also benefitted from their experience in carrying out educational evaluations. When, in 1965, Robert F. Kennedy and other federal legislators insisted that an evaluative element be included in the Elementary and Secondary Education Act (ESEA), a law that was to infuse massive federal dollars into the public schools, the necessity for widespread mandated educational program evaluation was established.

With the enactment of ESEA's evaluative stipulations, many educators wrote about what they thought educational evaluation should be. In these writings, numerous aspects of evaluation were considered. Definitions were created, models were developed, approaches were advocated, and methodologies were offered. During the last half of the 1960s, a diversity of educational writers contributed innovative notions regarding how educational evaluations should be conducted.

Because the field of educational evaluation was so new, the early writers had profound impact on how the field of program evaluation evolved. As already indicated, however, many of these writers profited from studying the efforts of others engaged in evaluation activities or from their own involvement in evaluation activities. Consequently, the thoughts and ideas proposed in earlier "experience-free" writings began to change. Early advocacies were reconsidered. Early certainties became more tentative. Early evaluative ideas have been, in many cases,

thoroughly restructured or significantly modified. Educational evaluation in the 1980s is not educational evaluation of the 1960s.

Purpose

This paper's mission is to consider the ways in which three evaluation theorists may have altered their views regarding two major evaluation-related issues during the past two decades.

Choosing three evaluation specialists for such a comparative analysis is not an easy task. After considering a number of evaluation theorists, it was decided that a review of the ideas of Michael Scriven, Robert Stake, and Michael Patton would provide interesting, yet informative insights regarding how, in at least two selected dimensions, the field of educational evaluation has changed in the last 20 years.

From Three Vantage Points

Scriven is known as the philosopher's evaluator, a powerful intellectual who sees the proper role of the evaluator as that of an "enlightened surrogate consumer" (Scriven, 1983). Evaluators, according to Scriven, should be equipped with (1) the skills to obtain pertinent and accurate information, (2) conversance with educationally relevant ethical considerations, and (3) an attentiveness to the common good (Scriven, 1967). Evaluators equipped with such skills should be focused on meeting the needs

of consumers. Presently, Scriven is a Professor of Education at the University of Western Australia.

Stake is noted for his significant and continuing contributions to educational evaluation. In 1967 his "Countenance Model"[1] was introduced. Stake's approach expanded on a previous Tylerian conception in which evaluators compare intended and observed outcomes. He did so by attending to the contextual variables and instructional transactions in relationship to standards and judgments (Stake, 1967). A 1975 paper substantially modified Stakes' earlier approach when he proposed the idea of "responsive evaluation." This updated approach called for continuing communication between the evaluator and audience in an effort to discover, investigate, and address the issues (Stake, 1975). Stake is currently a Professor of Education and the Director of the Center for Instruction and Curriculum Evaluation (CIRCE) at the University of Illinois.

Michael Patton, unlike Stake and Scriven, was not an early evaluation writer advocating a certain approach or methodology, but was an actual "in-the-trenches" educational evaluator. A sociologist by training, Patton's work has been influenced by Vito Perone, an experienced educational practitioner. Through Patton's studies of evaluation utilization, experiences in conducting evaluations, and his research in organizational dynamics and decision-making, he has come to his own views of the

---

[1]Stake prefers his "Countenance of Educational Evaluation" to be considered an "approach" or "persuasion" rather than being thought of as a model.

evaluation field in a unique and practical sense. His writings reveal that evaluation activities are not very useful, yet he contends that evaluation, above all, ought to be useful (Patton, 1978, 196). Patton serves as an international program specialist at the University of Minnesota.

Scriven, Stake, and Patton, then, will hopefully provide ideas that are philosophically-grounded in Scriven's case, responsively-oriented in Stake's case, and utilization-focused in Patton's case.

The Two Questions

As indicated earlier, two significant evaluation questions are to be considered in this analysis. The first question follows: "To what extent should results of educational evaluations be directed to single versus multiple audiences?" In the early days of educational evaluation, a fairly simplistic conception of educational evaluation was often advocated, namely, that evaluation studies were carried out to present decision-relevant evidence to a single decision maker. More recent thinking by those concerned with educational evaluation, however, suggests that evaluation results must be considered by a variety of relevant audiences. It would be potentially informative, then, to discover whether the views of three prominent educational evaluators have shifted or remained constant on this important question.

In the same vein, early conceptions of educational evaluation naively presumed the instant acceptance of an evaluation study's results and, therefore, the subsequent usefulness of those results to educational decision-makers. Experience, once more, has indicated that it is unrealistic to believe evaluation results will be immediately accepted, much less used. Thus, a second thrust of this analysis will be to investigate the degree to which educational evaluators should be attentive to the likelihood that results of their evaluation studies will actually be utilized by decision makers.

Following a restatement of each of the two major questions, the views of all three theorists will be presented using an historical approach. This isolation of early views and later views will be followed by a concluding analysis.

Question One: To What Extent Should Results of Educational Evaluations be Directed to Single Versus Multiple Audiences?

Scriven

A review of Scriven's early evaluation writings does not find the issue of single versus multiple audiences even addressed (Scriven 1967, 1972, 1974). A recent interview with Michael Scriven confirmed the issue's absence. Scriven said, "Originally I didn't give it (single versus multiple audiences) a great deal

of attention as I was looking at evaluation from a research-oriented point of view."[2]

Scriven indicated that with regard to this particular issue Robert Stake had substantially influenced his thinking. A review of the project format that Stake utilized for his evaluation study of the Illinois Project for the Gifted stimulated Scriven's own thinking on the use of multiple reports for multiple audiences.

Scriven's movement toward multiple audiences was further strengthened in 1977 when he was conducting an evaluation study. The initial findings of his evaluation were extremely negative and, if released, would have had serious adverse effects on the client. Putting the idea of multiple reports for multiple audiences to work, Scriven gave his client (in an executive session) a preliminary oral status report outlining the deficiencies he had thus far discovered. The client immediately implemented an action plan to ameliorate those deficiencies. Six months later Scriven conducted another evaluation from which a final report, outlining remaining program deficits, was written, submitted, and disseminated to various audiences.

Engaging in evaluation practice, then, has brought to Scriven's attention the need for a multiple-report strategy in order to supply information to more than a single audience.

---

[2]Scriven, Michael. Interview. Scriven's views are drawn chiefly from an interview conducted in Perth, Western Australia on March 24, 1987.

The idea of having multiple audiences appeals to Scriven but this multiple-audience notion, in his view, calls for rethinking how evaluation programs are conducted and how evaluators are trained. For example, if reaching multiple audiences is the main concern for an educational evaluation, then consideration of what evaluation approaches to utilize should be carefully considered.

Providing multiple audiences with information may require multiple evaluative approaches. In other words, instead of using one approach and then massaging the information to fit different audiences, one should plan various evaluation approaches to provide data for the various audiences from the beginning.

Evaluators, according to Scriven, should be trained with the multiple audience notion as an immediate reflex reaction. This reflex reaction can be accomplished by training evaluators to use a checklist technique that cues evaluators to look at a number of variables, but most importantly, process, outcome, and cost.

Scriven carries this checklist approach one step further by pointing out that a properly developed checklist not only cues the evaluator to obtain various kinds of information initially, but the checklist, upon completion, will lead to a series of multiple reports for multiple audiences.

This double checklist idea is represented in Scriven's (1987) The Double Checklist (KEC^2). Arranged in columnar-like fashion, the first column deals with "Evaluation 1" or, the evaluation of the program. The second column, "Evaluation 2", functions as an evaluation of the information acquired in

"Evaluation 1." (See attachment 1 for a copy of The Double Checklist (KEC_2),

Scriven expressed reservations, however, that the notion of multiple audiences is being corrupted by excessive responsiveness that may seem to distort truth. In other words, when using multiple reports the evaluator may be tempted to tell each individual group only what it is that it wants to hear instead of what the evaluator has actually uncovered. Scriven believes that an evaluator's first responsibility is to do a first-rate evaluation. The next responsibility is to communicate its results fully to all who should have access to the study's findings. As always, Scriven advocates that a carefully reasoned view of ethics should prevail.

Stake

Stake has always had, at the center of his evaluation thinking, the idea that evaluations should be focused, whenever applicable, on multiple audiences. Program evaluators should help audiences to see and improve what they are doing (Stake, 1967). This client-centered evaluation notion first appears in his 1967 "Countenance of Educational Evaluation."

In the 1967 article, Stake states that "an evaluation of a school program should portray the merit and fault perceived by well identified groups" (p. 112). According to Stake it is the evaluator's responsibility to "prepare a record of what educators intend," and "of what patrons generally expect" (p. 113). Thus,

the scope of services to the client in the 1967 model included meeting information-requirements as agreed upon by evaluator and clients at the outset of the study.

The notion of multiple audiences is further elaborated in his 1975 responsive evaluation essay. Stake wrote that responsive evaluation "orients more directly to program activities than to program intents, if it responds to audience requirements for information" (p. 292). Program evaluation, in Stake's reasoning, "pertains to the whole constellation of values held for the program and to the complex of expectations and criteria that different people have for such a program" (p. 291). Basically the scope of services in a responsive evaluation includes responding to audience requirements for information throughout the study.

As one can see, the scope of Stake's view of single versus multiple audiences has moved from an initially held narrow view of meeting prior agreements about information to a broader, continual responding to audience needs.

Stake contends that when attempting to respond to multiple audience requirements for information, a single collection of data will not suffice. Multiple audiences may require multiple data gathering which, of course, can be costly.

In order to avoid the increased costs, Stake recommends that evaluators try to be responsive to key issues, especially those held by people at the program site (Stake, 1975b). This responsive approach, however, takes more time than what Stake

refers to as "preordinate" evaluation approaches because evaluators must first become acquainted with the program and the evaluation context. Using the time necessary to get acquainted with the program, however, will allow the evaluators to utilize an adaptive approach for evaluation goal-setting and data-gathering requirements. Issues are suggested and discussed by the evaluation study's clients with the evaluators, resulting in the negotiation and selection of a few issues around which to organize the study. Once the limited number of issues are agreed upon, the appropriate data-gathering methodologies can be employed (Stake, 1975b).

A concern for multiple audiences, therefore, requires evaluators to realize that each issue to be studied may need additional or separate data gathering. Due to the contextual nature of evaluation, each evaluator must decide what to attend to in each situation.

Because Stake acknowledges that, in an attempt to be responsive, co-option may be a problem for evaluators, he advocates the use of meta-evaluation, a scheme whereby evaluations are themselves evaluated (Stake, 1975b). Meta-evaluation, a concept originally introduced by Scriven, can provide quality control measures for evaluation activities. Although Stake doesn't think evaluators deliberatively produce only favorable reports for clients, he does acknowledge that evaluators may be unconsciously, or even consciously, influenced by the clients.

Patton

Patton has always been of the opinion that results of educational evaluations should be directed to multiple audiences. His 1978 Utilization-Focused Evaluation outlines the first step to evaluation as being the "identification and organization of relevant decision-makers for and information users of the evaluation" (p. 61). The second edition of the same book addresses the same concern yet in more refined terms. Decision-makers and information users are now described as "the primary intended users of the evaluation" (p. 43). Patton contends that different people not only see things differently, but also have varying needs (Patton, 1978, 1986).

Identifying "the audience" for an evaluator is not enough. "The audience" can be a rather ambiguous group with heterogeneous interests and values. Instead, Patton seeks to identify the "primary stakeholders and intended users and why they are the key to the utilization process" (1986, p. 45).

When Patton was researching factors affecting the utilization of evaluation research, his data revealed what he characterized as a "personal factor." This "personal factor" was found to be present whenever evaluation findings were utilized. This "personal factor" has been defined as a "group of people who personally cared about the evaluation and the information it generated" (1986, p. 45).

Because Patton is concerned that the results of his evaluations be used, whenever he conducts an evaluation he

deliberately sets out to first identify the "personal factor," that is, to find and bring together people who want to learn, are enthusiastic, committed, competent, interested, aggressive, and, in his view, the primary intended users (Patton, 1986).

According to Patton, organizations do not consume information, but people do. Therefore, targeting evaluations at organizations is targeting them at nobody in particular (Patton, 1986). The "personal factor" is the essential ingredient of Patton's utilization-focused approach.

Patton's recent books (1980, 1981, 1982) also advocate the early isolation of relevant decision-makers and information users. Patton again affirmed his position regarding single versus multiple audiences in 1985 when attending an evaluation meeting in Malibu, California. He argued that evaluators should be aware that "different kinds of evidence speak to different people in different ways" (p. Malibu).

In sum, it appears that for Patton, early identification of multiple audiences, that is, primary intended users, has always been a given. He recommends that these primary intended users be identified in a carefully prescribed manner that focuses on determining and answering specific individuals questions.

Analysis

At first glance the evaluation positions of Scriven, Stake, and Patton appear to diverge sharply regarding the extent to which results of educational evaluations should be directed to

single versus multiple audiences. Closer scrutiny of the three positions, however, reveals that their current stances on this issue stem, to some degree, from mutual influences.

For example, although not discernible from the literature, it was the influence of Stake's work that initially brought the notion of audiences into Scriven's evaluation repertoire. Scriven, in turn, influenced Stake's early thinking as to how the term, "evaluation" should be defined (Stake, 1975b).

A reciprocal influence between Scriven and Stake is seen in the issue of co-option in evaluation. Co-option as a potential problem for responsive evaluations was recognized by Scriven and acknowledged by Stake. The solution to co-option, offered by Stake, was to use a Scriven-originated concept, meta-evaluation.

Patton, who entered the evaluative field almost a decade later than Scriven and Stake, has probably been influenced by both of their earlier writings. His more stable view, as opposed to their more shifting views, may be the result of his profiting from the experiences of writers such as Scriven and Stake. He does acknowledge their thinking and blends it into his own writings on evaluation (Patton, 1978, 1986).

Early writers in any specialized field tend to have more impact on one another because of the initial dearth of literature in early days. The field of educational evaluation has been no exception. The early writers, Scriven and Stake, seem to have influenced each other more than they have been influenced by later writers such as Patton. Patton's writings allude to the

views of Scriven and Stake. Yet, rarely is the influence reciprocated. Perhaps the more extensive body of evaluation writings available to Patton enabled him to develop views that are, by their practical and common sense nature, distinctive.

Whatever their origins or descriptions, the three theorists' positions seem to be relatively similar, with regard to aiming one's evaluation toward multiple audiences.

QUESTION TWO: To What Degree Should Educational Evaluations be Attentive to the Likelihood that the Results of their Evaluation Studies will Actually be Used by Decision-makers?

Scriven

A review of Scriven's early writings (Scriven 1967, 1972, 1974) finds, again, the absence of deliberation regarding the topic of evaluation utilization. Scriven, as revealed in a recent interview, said that he didn't discuss evaluation utilization in his early writings because he thought the crucial thing at the time was discussion of truth about the merit of programs. Because Scriven thought the focus of evaluations should center on the discovery of truth, he initially argued against Stufflebeam's CIPP model because the model was too decision-oriented.

Again, evaluation experience played a major role in influencing Scriven's thinking, prompting him to incorporate this more practically oriented belief into his already well conceptualized theoretical basis. To illustrate, when conducting a meta-evaluation for a project on urban studies, Scriven had the

opportunity to review 30 evaluations and discovered that very few of the evaluations were useful and, thus, that the needs of the client were not being met. Scriven acknowledges that this experience helped him understand why Stufflebeam was advocating such a decision-focused approach. Scriven now sees a need for program evaluations to focus their activities on producing information for actual decisions.

In the early 1970s as Scriven came to think about the issue of evaluation utilization, he made concessions that evaluation reports should be written in a language the client can understand, in terms of the decisions a client faces, in terms of the needs a client has, and in terms of the needs the consumer has. Practical field experience, then, increasingly broadened his evaluative views.

There are, however, two potential troublespots for utilization-focused evaluation that concern Scriven. First, making an evaluation useful to decision makers is legitimate as long as evaluators submit truthful reports to clients. Scriven admonishes Cronbach for taking the position that evaluators shouldn't submit negative findings to decision makers because decision makers are unlikely to accept such findings. Not submitting a report on the basis of the report being damaging to your client (who is paying you) is unethical. In other words, heightened attention to making evaluation reports relevant to the decision maker is important, but never to the point of withholding negative information even if it would give "bad

press" to the client. Secondly, the idea of not submitting an evaluation report that contains recommendations that are not within the client's power to fulfill is also unethical in Scriven's view. Above all, Scriven subscribes to the importance of truthful reporting and complete disclosure. In sum, evaluation utilization reporting must be ethical, containing honest and viable recommendations.

Stake

Stake has always been concerned that results of evaluation studies be utilized by decision-makers. Prior to publishing his "Countenance of Educational Evaluation," Stake had begun to sense the futility of the current status of educational evaluation when he wrote, "Today's educator may rely little on formal evaluation because its answers have seldom been answers to questions he is asking" (p. 107). The 1967 article was written after Stake witnessed Cronbach and Scriven argue about the relative needs of producers and consumers and it became apparent that "something was needed in the way of an overview matrix of the different data that could be used" in evaluation (p. 92).

Stake, realizing that his original article did not address the idea of how to process data, wrote a 1975 article where he described responsive evaluation as being a naturalistic, phenomenological, or case-study approach to evaluation where there is more of a political than methodological commitment (Stake, 1981). Responsive evaluation involves identifying

certain people (stakeholders, as others have defined them) "who have an important concern about the program and who should be considered carefully in deciding what the design of an evaluation study should be" (p. 92). Thus, as evidenced through his writings, Stake sees that the potential for evaluation utilization is accomplished by evaluators responding to the needs of the people, by allowing the views and needs of the stakeholders to shape evaluation purposes.

More recently, Stake has again demonstrated his concern for evaluation utilization. In a 1980 paper, Stake discusses what he sees as the utilities and disutilities of evaluation. He begins his discussion by pointing out that "Formal evaluation studies of educational programs are undertaken with the expectation that they will be useful. The investigators, the sponsors, the program staff and various other stakeholders have different uses to which they may put the studies. The allocation of resources of investigation may vary considerably as different potential uses are considered." When speaking of the utility of evaluation, Stake observes that "Decision making and policy setting are important consequences of evaluation study." But Stake also points out that whether or not decision making or policy setting occurs as a result of evaluation, "<u>understanding</u> alone is a consequence that indicates utility of evaluation."

Moreover, it is important, according to Stake, to take into account that the "<u>imprimatur</u> of evaluation studies...can be itself a substantial utility." In other words, actually

evaluating a program can enable (1) a program to be viewed as more authentic or (2) those who authorized the evaluation to be viewed as acting in a responsible way.

There are, in Stake's view, disutilities in evaluation as well. Conducting evaluations may interfere with or postpone program functioning. Conducting evaluations costs money that may be diverting dollars from other projects. Even worse, the knowledge gained in evaluation may serve to "obscure key understandings or impede certain decisions."

For Stake, then, utility in evaluation can arise from (1) influencing decision making and policy setting, (2) enhancing understanding of a program, or (3) serving as an imprimatur for a program.

Stake, like Scriven, addressed the issue of truthfulness in utilization-focused evaluation reports. In a 1982 essay, "How Sharp Should Be the Evaluator's Teeth?" Stake contends that "sometimes instructive, sometimes demanding, the evaluator should know how to bite gently." Stake favors "strong recognition of the social role of evaluation when deciding how much to probe and expose the frailty of the program and even its iniquities." Clearly, Stake desires truthfulness in evaluation reporting, yet recognizes the potential for utilization that exists if truth is combined with tact.

In sum, Stake sees responding to the needs of those people who have an important concern about the program, i.e., the

stakeholders, as the method for accomplishing the goal of evaluation utilization.

Patton

Utilization of results by decision makers has always been the central goal in Patton's evaluation work. Although the literature on program evaluation never explicitly defines utilization, Patton sees that there is an implicit definition: "Utilization occurs where there is an immediate, concrete, and observable effect on specific decisions and program activities resulting directly from evaluation findings" (p. 30). Committed to narrowing the gap between generating evaluation information and actually applying evaluation information for decision making, Patton has organized his whole approach around this utilization-focused theme.

Because Patton's focus in evaluation is on intended use for intended users, he believes that this orientation leads to major implications that must be considered by those who would use his approach. First, evaluators must find and bring together the primary intended users, that is, people who want to know something and are able and willing to use information. Care must be taken when identifying primary users to find strategically located people. Patton has discovered that lower level people in the organizational hierarchy, as opposed to those in authority by virtue of their formal positions, are much more likely to employ evaluation findings.

Next, evaluators must identify and focus the evaluation concerns which directly answer the questions of the primary stakeholders. Evaluative methods that generate this useful information must be determined. It is Patton's contention that this shared decision-making process distinguishes utilization-focused evaluation from other evaluative approaches.

Finally, the primary users must be consulted in data analysis, data interpretation, and data dissemination. Clearly, then, being a utilization-focused evaluator requires considerable energy. In Patton's opinion, the payoff, in terms of program improvement, is worth the effort (Patton, 1986).

In addition to the three major implications, Patton sees two main factors that affect the utilization of evaluation results. First, the "personal factor," which was discussed earlier in this paper. Second, is the "political factor." According to Patton, a basic assumption underlying utilization-focused evaluation is that the program evaluation is "a highly political process from initial conceptualization to ultimate utilization" (1978, p. 40). Both his 1978 and 1986 utilization-focused evaluation books have devoted considerable attention to describing the political nature of evaluation. This political process is naturally created by the various stakeholders' advocacy of certain values and perceptions. As Stake has previously asserted, Patton also argues that evaluators must be sensitive to and respectful of varied interests. In sum, evaluators who desire success must be alert to the realities of the political process.

In achieving evaluation results that are utilization-focused. Patton, again, like Stake, advocates the use of qualitative approaches whenever needed (Patton, 1980). Unlike data derived from strictly quantitative methods, qualitative data provide depth and detail about the evaluation setting. If evaluation research is to be useful, then different methodologies, e.g., qualitative approaches that are creative and practical, may be necessary in order to obtain useful information for decision-making (Patton 1981, 1982).

Achieving the aim of evaluation utilization, then, involves substantial effort on the part of evaluators. A high level of skill and knowledge must be employed in order to make utilization the driving force in an evaluation. The concern for utilization must begin when the evaluation is first conceptualized, continuing throughout. The evaluation should be aimed at the interests and information needs of the identified stakeholders. The intended evaluation user should be actively involved throughout the evaluative process in order that the evaluation data produced are both meaningful and relevant.

Analysis

When reviewing the positions taken by Scriven, Stake, and Patton concerning the degree to which educational evaluators should be attentive to the utilization of evaluation findings, it is evident that, albeit to a lesser degree, mutual influence is

again apparent along with a general concurrence that evaluators should be attentive to the utilization of evaluation findings.

There is a growing awareness and growing acceptance on the part of Scriven and Stake regarding this issue. But their awareness and acceptance lags far behind Patton's plan for achieving the utilization of evaluation results.

Scriven wants to have evaluation results used, but he sees utilization as clearly secondary to the evaluation's being a first-rate assessment of merit. Truth and discovery, rather than information and utilization, are the heart of Scriven's evaluation creed.

Stake sees that the way to ensure evaluation utilization is by making evaluations responsive to the needs of the people. Using a broad notion of utility, he certainly acknowledges the importance of the issue. Yet, it may represent wishful thinking when Stake argues that responsive evaluation will, in fact, be used.

Patton, however, is more than just attentive to the likelihood that evaluation findings will be used. He makes evaluation-utilization the mainstay of his approach in an effort to ensure utilization.

Epilogue

Does a convergence or divergence of thought, over time, exist in the thinking of Scriven, Stake, and Patton on these two selected evaluation dimensions? Surprisingly, the philosopher-

evaluator, the responsive-evaluator, and the utilization-focused evaluator start from very different vantage points, but through time and, as a result of experience, their separate positions have begun to move toward the same goal. Their views concerning attention to multiple audiences and evaluation utilization are slowly moving in the same direction. The potentiality for convergence appears to be a possibility.

REFERENCES

Patton, M. (1978). Utilization-focused evaluation. Beverly Hills, CA: Sage

Patton, M. (1980). Qualitative evaluation methods. Beverly Hills, CA: Sage.

Patton, M. (1981). Creative evaluation. Beverly Hills, CA: Sage.

Patton, M. (1982). Practical evaluation. Beverly Hills, CA: Sage.

Patton, M. (1985). Evaluation reflections: Malibu Meeting Book, (M. Alkin, ed.), Center for Study of Evaluation, UCLA.

Patton, M. (1986). Utilization-focused evaluation. (2nd edition) Beverly Hills, CA: Sage.

Scriven, M. (1967). The methodology of evaluation. In Perspectives on Curriculum Evaluation, (AERA) Monograph Series on Curriculum Evaluation, No. 1. Chicago: Rand McNally.

Scriven, M. (1972). Prose and cons about goal-free evaluation. Journal of Educational Evaluation, 1972. 3, 1-5.

Scriven, M. (1974). The concept of evaluation. In M.W. Apple, M.J. Subkoviak, & H.S. Lufler, Jr., (eds.), Educational Evaluation: Analysis and Responsibility. Berkeley, CA: McCutcheon, 1974.

Scriven, M. (1975). Evaluation bias and its control. Occasional Paper Series No. 4, Western Michigan University, Evaluation Center.

Scriven, M. (1983). Evaluation ideologies. In G.F. Madaus, M. Scriven, & D.L. Stufflebeam (Eds.), Evaluation models. (pp. 229-260). Boston: Kluwer-Nijhoff Publishing.

Scriven, M. (1987). The Double Checklist (KEC). (unpublished)

Scriven, M. (1987). Interview, March 24, 1987, Perth, Western Australia.

Stake, R. (1967). The countenance of educational evaluation. Teachers College Record, 68, 523-540.

Stake, R. (1975b). Program Evaluation: Particularly responsive evaluation. Kalamazoo: Evaluation Center, Western Michigan University, Occasional Paper No. 5, November.

Stake, R. (1980). "The Utility of Evaluation" (unpublished)

Stake, R. (1981). Interview. Educational Evaluation Policy Analysis. Vol. 3, No. 3, p. 91-94.

Stake, R. (1982a). Persuasions, Not Models. Evaluation Research Society Newsletter, Vol. 6, No. 2, Summer, 1982, p. 8.

Stake, R. (1982b) "How Sharp Should be the Evaluator's Teeth?" (unpublished)

THE DOUBLE CHECKLIST (KEC_2)

	EVALUATION 1 (of the program)	EVALUATION 2 (of evaluation 1)
DESCRIPTION	Correct, not official one. For overview & identification*, not installation. Identify components* if any, and their functions, if clear. (*terms are defined on the following page)	Full design and staffing; cover all data needed for checkpoints, fix all specifications/parameters; use metaevaluator* here.
BACKGROUND	History, relation to other programs, precedents, enemies, allies, connections. Who gave you descriptions or evaluations that turned out to be completely wrong/accurate?	Same; who demanded or requested the evaluation, were there prior evaluations etc.?
CONSUMER*	Targeted/impacted/potential. Recipients*/providers*/stakeholders*	Client*/audiences* Prior evaluations etc.
RESOURCES	'Strengths Assessment' Defines the potential	Constraints on design
VALUES	Needs/Wants/Standards/Ideals Legal/Scientific/Ethical/Hedonic/Professional/Aesthetic	Here we look at timeliness, relevance, cost-effectiveness etc.
PROCESS	a. Process description recycles to Description b. Process evaluation using Values and valid indicators. For example; Is promotion based on merit? Did the program respect confidentiality?... Has the program set up to facilitate evaluation....	Might use 'implementation index' and did the evaluation do so too? and was the evaluation, too?
OUTCOMES	Intended/Incidental/Unexpected (does not include expenditures)	Includes report(s) and its effect
GENERALIZABILITY	Across sites, time, staff, etc. Especially for the future; will the needs be the same, will the enthusiasm be there?	Other uses for design, later, for different programs etc.
COSTS	Money/non-money x start-up/support/opportunity	

COMPARISONS Identify the critical competitors*, which should include significantly cheaper, significantly better, and radically different; also, may include some possible as well as actually available alternatives. Normally will include those the decision makers have in mind, but rarely only those.

SIGNIFICANCE	Comparative cost-effectiveness etc. The synthesis of all the above	Use a metaevaluator for this AND the client/audiences.

RECOMMENDATIONS or **policy implications** (optional) These normally require extensive area expertise.

Description The description should contain little evaluative language, and what there is must be non-controversial; you can say that the equipment is new, but to say it is in good condition will need a little more support, and to say that it is ideal will have to be shown as one of the conclusion of the evaluation, not a premise. The description must be good enough so that one could be sure that one has replicated the program at another site; but it's not part of the evaluation (usually) to go into all the details of how to set up (install) the program at another site.

Components Examples include: the delivery system, the support system (or 'infrastructure') – e.g. administrators, office staff, training staff, maintenance staff, internal evaluation team, consultants – the equipment (vehicles, computers, copiers etc.), the buildings, the environment.

Metaevaluator Usually a distinguished evaluator who can critique the design and execution of the evaluation (formative and summative evaluation). The evaluator should get a copy of the summative report in just the way that the program staff gets a copy of the program evaluation i.e. with a chance to reply and the guarantee that the report will be changed until it satisfies the evaluand OR the objections will be forwarded to the client at the same time as the report. The m/evaluator should be sufficiently senior and/or busy to be relatively immune to the lure of future income from repeats of this role, if paid, by comparison with the professional need to do good work.

Consumer Anyone affected by the program, directly or indirectly, intentionally or unintentionally; this always includes the taxpayers when we are dealing with publicly funded programs

Recipients are the immediate consumers; other consumers are affected indirectly ('ripple effects')

Stakeholders are those who have some ego or other capital invested in the program, and can be held to be in some way responsible for it, but may not be actively involved in the day to day operations; e.g. inventors or instigators or supporters of the program

Audiences Those who will read or hear of the evaluation, either during or at the end of the process, including many who are and many who are not being evaluated.

Providers are mainly the staff who are serving the immediate consumers; they are also stakeholders

Client is the person/group who has commissioned the evaluation; the consumer of the evaluation. The term is also often used for recipients of social services such as counseling that involve some client-professional relationship. It's sometimes easier to use the term 'recipient' for the latter group (one of the consumers), when discussing an evaluation, to avoid confusion

Critical Competitors are the alternative options that will produce comparable effects, as seen from the point of view of the consumers. They may not look like the evaluand at all; may have to be created especially for the exercise; may not have been thought of before the evaluation began. Thus they are (more or less) <u>functionally equivalent</u>, but not necessarily similar in terms of Description or Process